Title: Inferring motives in psychology and psychoanalysis

Author: Michael Lacewing

Contact:

Heythrop College

Kensington Square

London W8 5HN

m.lacewing@heythrop.ac.uk

020 7795 4223

Biography:

Michael Lacewing is Director of Research and Senior Lecturer in Philosophy at Heythrop College, London. He researches in philosophy of psychoanalysis and moral psychology, is co-editor (with Louise Braddock) of *The Academic Face of Psychoanalysis* (London: Routledge 2007), and has recent publications in *Ratio* and *Philosophical Psychology*.

Word count: 9917

Character count (incl. spaces): 64,154

(Microsoft Word for Mac.)

Inferring motives in psychology and psychoanalysis

Abstract

Grünbaum (1984) argues that psychoanalysis cannot justify its inferences regarding motives using its own methodology, as only the employment of Mill's canons can justify causal inferences (which inferences to motives are). I consider an argument offered by Hopkins (1988) regarding the nature and status of our everyday inferences from other people's behavior to their motives that seeks to rebut Grünbaum's charge by defending a form of inference to the best explanation that makes use of connections in intentional content between behavior and motives. I argue that Hopkins succeeds in defeating Grünbaum's objection as it is presented, but that work in social psychology presents a further challenge. I discuss the extent to which the challenge can be met, and conclude that certain types of inference in psychoanalysis are justifiable, but others, including those which are the target of Grünbaum's objection, cannot be justified by the methods defended by Hopkins.

Keywords:

Commonsense psychology, scientific psychology, inference to the best explanation, Adolf Grünbaum, James Hopkins

Inferring motives in psychology and psychoanalysis

In this paper, I consider an argument offered by Hopkins (1988) regarding the nature and status of our everyday inferences from other people's behavior to their motives and other mental states. (It may be that we recognize, rather than infer, immediate intentions and emotional states, e.g. in bodily actions and facial expressions. But I am concerned with inferences that go beyond states that can be recognized immediately in this way.) Hopkins' argument seeks to rebut the charge, leveled by Grünbaum, that psychoanalysis cannot justify its inferences regarding motives using its own (clinical) methodology. Hopkins and Grünbaum agree that inferences to motives are causal inferences. Grünbaum argues that only the employment of Mill's canons can justify causal inferences, while Hopkins defends a form of inference to the best explanation that makes use of connections in intentional content between behavior and motives. Hopkins argues that the clinical methodology of psychoanalysis is an extension of that of commonsense psychology, so that if Grünbaum's objection to psychoanalysis were cogent, it would equally undermine our everyday psychological inferences. I will argue that Hopkins succeeds in defeating Grünbaum's objection as it is presented, but that work in social psychology on our everyday inferences regarding the causes of behavior presents a further challenge. I discuss the extent to which the challenge can be met and the implications for psychoanalysis.

§1. Grünbaum's challenge to psychoanalysis

§1.1 The challenge

Grünbaum (1984) critically discusses the methodology by which, he claims, psychoanalysis developed its claims regarding the causal origins of neurosis. Clinical data comprise the totality of patients' verbal and non-verbal behavior in the psychoanalytic clinical setting; this behavior is

(ideally) generated by the patients' free associations, beginning with whatever topic first enters their mind at the start of the session, and includes their free associated responses to the analyst's interpretations as the session unfolds. From this data, Grünbaum argues (1984, 6-7), including the evidence of therapeutic success, Freud inferred that neurosis had its origin in the repression of an experience or memory, and the same methodology continues to be used by many contemporary psychoanalysts to justify developments in the theory of neurosis. However, psychoanalytic theory cannot be justified by this method for two reasons. First, because clinical data may be contaminated by suggestion, causal hypotheses resting on clinical data are evidentially unsupported. Second, the methodology employed for inferring causes from clinical data is flawed. Grünbaum concludes (1984, 265-6) that psychoanalytic claims regarding the causes of neurosis require confirmation (or refutation) by extraclinical studies (he does not claim that the psychoanalytic theory of the mind is untrue).

I shall not discuss Grünbaum's first argument here. Setting aside the issue of suggestion is not unfair to Grünbaum, as he states (1984, 128) that even if we knew clinical data were uncontaminated, psychoanalysis still could not validate its causal inferences, and it is this claim with which I am concerned.

It is a common response to Grünbaum that he attacks a set of claims in psychoanalysis that have long been superseded, as he focuses on Freud's early theories regarding neurosis. Psychoanalysts no longer hold that neurosis arises in law-like fashion from the repression of a (set of) childhood experience(s). The clinical evidence does not support such a view, and many psychoanalysts today would support a characterological account of neurosis, along the lines of 'neurosis is the result of chronic psychic conflict together with the failure of psychic defense'. But Grünbaum's challenge is whether psychoanalysis can validate its causal claims about neurosis by its own methods, and the challenge is intended to hold even as one theory about neurosis is

replaced by another. The view that contemporary psychoanalytic theories are better supported by the clinical evidence will only be acceptable once we have clarified and defended the method by which clinical data can support any such causal claims. My argument will therefore focus on methodological concerns alone.

The principle of free association remains fundamental to psychoanalysis. Psychoanalysts claim that the interpretation of free associations can discover the psychological states that are the unconscious causes of the clinical data – the verbalizations, conscious emotions and transference phenomena. But why think that free association from some 'effect', e.g. a neurotic symptom or action, dream, or slip, will lead towards its cause (the unconscious material), enabling the cause to be inferred by interpretation of the clinical data? Grünbaum argues that Freud's original argument was this: the patient free associates from a neurotic symptom; the psychoanalyst interprets; where the interpretation (together with the process of 'working through') lifts a repression, the symptom is removed; repression therefore causes neurotic symptoms; free association therefore leads from effect to cause. However, the causal inference from lifting repressions to repression causing neurosis is flawed. It is possible that some other cause, e.g. suggestion, brought about symptom removal. To rule out this possibility, it needs to be shown that lifting repressions is uniquely (or at least significantly) causally effective in symptom removal. Even then, it cannot be inferred that repression caused neurosis in the first place, only that repression is necessary for symptoms to be maintained. If the removal of symptoms does not depend on discovering the (repressed) causes of the symptom, we cannot infer that what has been discovered through free association to a symptom is a cause of that symptom (and likewise for dreams, slips, etc.). The causal relevance of the 'end' of the chain of free association to the 'beginning' of that chain has yet to be demonstrated.

§1.2 Grünbaum on causal inference

With the failure of the argument from therapeutic success, Grünbaum argues, we must employ the standard method for inferring causes:

> A causal hypothesis of the sort encountered in psychoanalysis asserts that some factor *X* is *causally relevant* to some occurrence *Y*. This means that *X makes a certain kind of difference* to the occurrence of *Y* in some reference class *C*…
>
> To validate a claim of causal relevance, we must first divide the reference class *C* into two subclasses, the *X*'s and the non-*X*'s. And then we must show that the incidence of *Y*'s among the *X*'s is *different* from what it is among the non-*X*'s. But it is of cardinal importance to appreciate that *this requirement is entirely neutral as to the field of knowledge or subject matter*… (2004, 146; italics in original)

An oversimplified example: to establish whether smoking (X) is causally relevant to lung cancer (Y), we must compare the incidence of lung cancer in smokers and non-smokers; and/or we must examine cases of lung cancer to see whether the person smoked or not. If there is a higher probability that a smoker will have lung cancer than a non-smoker, we may infer that smoking causes lung cancer. Likewise if there is a higher probability that someone with lung cancer is a smoker than non-smoker.

Grünbaum is appealing to Mill's (1904) canons, of which the Methods of Agreement and Difference are of most importance here. We employ the Method of Agreement when, finding only one antecedent shared by all instances of some effect, we infer that to be a cause. We employ the Method of Difference 'when we find that there is only one prior difference between a situation where the effect occurs and an otherwise similar situation where it does not,

[and] infer that the antecedent that is only present in the case of the effect is a cause' (Lipton 2004, 18). According to these canons, to establish that *X* is the cause of *Y*, we must show that *X*s make a difference to the incidence of *Y*s *by comparing (classes of) cases*, either comparing the incidence of *Y*s in cases in which *X* occurs with ones in which it does not (Difference) and/or by comparing cases in which *Y* occurs to see whether *X* occurred or not (Agreement).

On the face of it, inferring that the cause of neurosis is the repression of a memory on the basis of clinical evidence does not meet the standards of evidence of Mill's canons. On the assumptions that there are people who are not neurotic, and that people who are not neurotic do not seek psychoanalytic help, clinical data tells us nothing about people who are not neurotic. Therefore, it cannot support the claim that the difference between neurosis and its absence is the occurrence of repression (or chronic psychic conflict together with the failure of psychic defense). Furthermore, we need to use Mill's canons to establish a causal link between the clinical data produced through free association and the inferred unconscious psychological states. We cannot discover the causes of dreams, slips or neurotic symptoms and actions by interpreting the clinical data alone.

A variety of responses are open to the defender of the clinical methodology of psychoanalysis:

1. reject Grünbaum's claim that psychoanalysis makes causal inferences (the 'hermeneutic' interpretation);
2. reject Grünbaum's characterization of the clinical method, e.g. arguing that psychoanalysts have in fact employed Mill's canons, albeit implicitly;

3. reject Grünbaum's commitment to Mill's canons, arguing that psychoanalytic theorization conforms to the standard scientific method for justifying causal claims (e.g. a defense of inference to the best explanation in the philosophy of science);
4. reject Grünbaum's claim that causal inferences must *always* be justified on the basis of Mill's canons, arguing for the validity of another method, which is, in fact, employed in psychoanalytic inferences from clinical data.

My primary focus is Hopkins' version of (4). (2) shall arise during the discussion in §3.1. Miller (1988) presents a version of (3), but consideration of this would take us too far afield, into general topics in philosophy of science. Furthermore, it does not avoid the challenge presented in §4 (see note 2). I shall not discuss (1), the hermeneutic interpretation. The arguments in its favor claim either that psychoanalysis, specifically, does not deal in causes but meanings; or more commonly, that any account of human behavior that seeks to 'make sense' of it is non-causal (a view, embodied in many variations, whose slogan might be 'reasons are not causes'). Against the first, I believe Grünbaum has demonstrated that psychoanalysis is best understood as seeking a true account of the underlying psychological causes of behavior. Against the second, while I cannot here defend the (widely-held) view that inferences regarding motives are causal inferences, some brief comments are in order to situate the debate and its commitments. The 'standard' explanation of action, deriving from Davidson's famous 1963 paper, 'Actions, Reasons, Causes', maintains that an agent's beliefs + desires afford the agent reasons, and reasons are causes. This standard model has recently come under heavy critical fire (see, e.g., Sandis 2008). It can be argued that reasons are not psychological states, but their content; and that psychological states are not causes of actions, as it is agents, not their beliefs and desires, that bring about actions. However, one does not have to accept the standard model to accept

that explanations of actions in terms of motives are causal explanations, and that inferences to motives are therefore causal inferences, e.g. if one accepts that motives are causally relevant to (rather than causes of) actions. It is sufficient for my purposes to be able to say that an agent acted as she did in virtue of certain motives, and for this statement to be understood causally. It is standard in causal explanation to mention not just the cause, but those features of the cause in virtue of which it brought about its effect. It is also worth noting that this position is not committed to the view that all causally relevant factors operate according to laws. Mental states, such as beliefs, desires, wishes, phantasies, memories, and the like, make a difference to how we behave, though there may be no strict laws that govern the relations between such states and their effects.

§ 2. The commonsense psychology defense of psychoanalysis

§ 2.1 The 'method of interpretation': Psychoanalytic inferences as inference to the best explanation

In large part, psychoanalysis is a psychology of motives, construed broadly to include desires, wishes, drives, impulses, emotions and other affective phenomena. So is there a methodology for generating and justifying causal inferences regarding motives that is independent of Mill's canons? Hopkins argues that there is, and his defense of psychoanalytic causal inferences rests on two key assertions:

> In commonsense psychological practice we already establish causal connections (in particular concerning the role of motives) interpretively, in ways that are autonomous, cogent, and prior to such canons. So it seems wrong to hold generally that cogency in a psychology of motive must satisfy them…

> Further, psychoanalytic theory seems an extension of commonsense understanding of motives, by interpretive means internal to it. (1988, 37)

Premise 1: Hopkins argues that we interpret voluntary behavior on the basis of its apparent intentional content, e.g. as 'getting a glass of water', 'studying hard', 'seeking revenge'. In so doing, we place it in relation to a motive or set of motives, which we take to be causally relevant to the behavior. In the first instance, this is the desire or intention with the same intentional content as the behavior, e.g. the desire or intention to get a glass of water, study hard, or seek revenge. The motive both explains and is causally relevant to the behavior in virtue of its intentional content, while the behavior 'inherits' its intentional content from the causally relevant motive.

Most of these causal inferences are not possible, or at least justifiable, on the basis of a single example of behavior, with no background information. We cannot know that someone is seeking *revenge* without knowing more about their beliefs about the previous actions of those they seek to harm. That information may be gleaned from other behavior, including verbal behavior. The interpretation that someone is seeking revenge therefore rests on a range of behavior, and, importantly, integrates it into a pattern. The connections of intentional content and of causal relevance that hold between motives and actions we take also to hold between mental states, so that motives (revenge) may themselves be motivated (by anger at a harm done). The explanatory patterns involve patterns of similarity and connection in the intentional content of the mental states and behavior: the act of *seeking revenge*, the desire and intention *to seek revenge*, where the ground of *revenge* is *a previous harm or offence inflicted*, the belief that some prior act *inflicted harm or offence*, perhaps the evaluation that the act *merits revenge* or more simply, the standing disposition *to avenge harm or offence*, and so on.

If an explanation we initially entertain does not fit into a broader pattern of motivation for which we already have evidence, we revise the explanation and/or the pattern. It takes no imagination to supply the kinds of patterns which would lead us to reject an inference that someone is drinking because they are thirsty:

a. the behavior does not fit a pattern of voluntary action, or
b. another pattern indicates the agent is not thirsty (e.g. avowals to this end, having just drunk a considerable quantity of water and not suffering from dropsy…), or
c. other evidence indicates an ulterior motive could be at play (e.g. the drinking follows a challenge to see how much the agent can drink; the agent has the strong belief that they should drink eight glasses of water a day, irrespective of thirst…).

It is clear from this example that inferences from action to motive must fit not only with other such inferences, but also with background knowledge about the operation of (this and other) motives. The reasoning we employ is a form of inference to the best explanation, and the virtues we aim at, and in the light of which we revise our inferences, are the well-known ones of accuracy (fit with evidence), scope, coherence/consilience, simplicity, and plausibility (fit with background knowledge). I shall call this the method of interpretation. (For those with an interest in these matters, nothing in the method of interpretation commits one to a particular viewpoint in the theory theory v. simulation theory debate, as both positions agree that we work with intentional content and utilize background knowledge, and are compatible with the view that we employ inference to the best explanation.)

Premise 2: Hopkins argues that inferences in psychoanalysis, based on free association, operate in the same way. The place from which free associations begin (symptom, dream, slip)

has a certain intentional content, though its connections to any motives are at the start obscure. Free association provides further material, rich with intentional content, from which an interpretation regarding the intentional content of the motive may be made. Just as we infer from action to motive, we infer from starting point + free association to motive. Just as such inferences in the case of commonsense psychology become both more justified and more explanatorily powerful as more evidence is accrued, more behavior explained, so are inferences in psychoanalysis, resting on the clinical data produced by the analysand over many sessions. (Just such an argument appears in Freud (1937).) Psychoanalytic claims regarding the causes of neurosis are justified (or not) by the standards of inference to the best explanation, operating in a way similar to the way it operates in commonsense psychology.

Now, it may be that specific psychoanalytic hypotheses regarding the causes of neurosis are *not* supported by the clinical evidence, using the method of interpretation. But, as emphasized at the outset, Grünbaum's objection is one of method, not truth, and Hopkins has suggested an alternative method by which causal inferences in a psychology of motive may be justified.

It has been contested, e.g. by Manson (2003) and Kitcher (1992), whether the similarity between causal inference in psychoanalysis (at least, as Freud practiced it) and commonsense psychology is as great as Hopkins makes out, in particular, whether psychoanalysis draws on theories that do not originate in commonsense psychology. Hopkins argues the point at length, using examples from Freud, and a recent piece on contemporary psychoanalytic methodology by the psychoanalyst Michael Brearley (2008) supports Hopkins' analysis. As only detailed consideration of examples of psychoanalytic reasoning can settle this matter, I shall assume for the purposes of my argument that Hopkins is essentially correct in his account of psychoanalytic inference to concentrate on the question of whether such inference is justifiable.

§ 2.2 Millian canons in psychological inferences

Grünbaum explicitly rejects the form of causal inference defended by Hopkins, dubbing it the 'thematic affinity fallacy'. 'Thematic affinity' picks out the similarities of intentional content that hold between action (or symptom/dream/slip) and motive, and Grünbaum denies that we can ever legitimately use such a ground for inferring causal connections. Suppose someone who desires to read a certain book, and believes the book is in the library, visits the library. His desire qualifies as causally relevant to his visiting the library, Grünbaum says, 'by affecting the incidence of visits to a library' (1984, 73). This is the phraseology of Mill's canons, and so Grünbaum appears to be recommending their employment here. We cannot know from a single instance, or even a statistically insignificant sample of instances, that the desire to read a book that one believes is in the library affects the *incidence of visits* to a library. Instead, employing the Method of Agreement, we need to examine the class of visits to a library (VL) and compare it with the class ~VL. If we find that the (unique or primary) distinguishing factor between the two classes, is the presence of a desire to read a book that the subject believes is in the library, and that the difference between the classes is statistically significant, we may infer that the desire is a cause of such visits – it has affected the incidence of such visits. Employing the Method of Difference, we need to compare the class of desires to read a book that the subject believes is in the library (RB) with the class ~RB. If we find that the incidence of visits to the library is greater in RB than ~RB, and the difference is statistically significant, then we may conclude that such a desire is causally relevant to visits to the library.

That this is what Grünbaum has in mind is indicated by a second example:

> I can validate by means of these [Mill's] methods a supposed causal link between my conscious awareness of a derogatory remark and my introspectively experienced anger by noting, over a period of time and repeatedly, that the presence vs. the absence of insults *makes a difference* statistically to my becoming angry. (1993, 164)

However, we do not naturally take ourselves to be inferring motives on such grounds, and it would take an argument to convince us that this type of statistical reasoning is what we are doing, despite appearances to the contrary. If we do not reason like this, Grünbaum's insistence that Mill's canons are the *only* way to establish causal inferences would 'render all everyday understanding [of motives] groundless' (Hopkins 1988, 52).[1]

It is noteworthy that 'scientific' psychology, as opposed to commonsense psychology, commonly adopts Grünbaum's approach to inferring causes. If Hopkins is right about commonsense psychology, there is a methodological gap between commonsense psychology and scientific psychology, and psychoanalysis falls on the side of commonsense psychology – which may help explain its exclusion from scientific psychology. In this debate, therefore, there are significant issues at stake regarding psychology's self-understanding.

§3. Causal inference in commonsense psychology

We need to examine more closely Grünbaum's criticism of causal inference by 'thematic affinity' and consider whether his objections apply to Hopkins' method of interpretation. In doing so, I consider the suggestion, raised in the second response to Grünbaum from §1.2, that we (and psychoanalysts) may be employing Mill's canons implicitly, and it is this that justifies our causal inferences.

§3.1 Background knowledge and non-conscious statistical tracking

Grünbaum accepts that many of our causal inferences in commonsense psychology are justifiable, while his examples indicate his belief that we employ Mill's canons. However, the examples suggest that we employ Mill's canons implicitly in relation to each causal inference we make, in turn. This, I think, is a mistake, and unnecessary. It is a mistake because we are happy to make inferences to motives that we have not encountered before, or in situations that are unfamiliar. Given the fact that we cannot have implicitly tested the causal pattern of the motive (the relative incidence of its effects) in relation to the situation, either any such inference is unjustified or we are guided by background knowledge. Given the skeptical conclusion of the first alternative, the second is more appealing. And it is unnecessary as Grünbaum can embrace this result, given his account of the general role of background knowledge in justifying causal inferences.

In a two-page discussion (1984, 259-60), Grünbaum considers exceptions to the use of Millian canons in causal inference. Astronomers were justified in making causal inferences regarding planetary motion without experimental controls by drawing on their background knowledge of Newton's laws of motion – which had been experimentally demonstrated. Again, we are justified in inferring from a footprint in the sand that it was caused by a foot. The inference is valid, not on the basis of the similarity ('thematic affinity') between the shape of the print and of a foot alone (though this is relevant), but on the basis of background knowledge that such shapes are not caused by chance, e.g. through the action of the wind (2004, 152-3). Causal inference, therefore, is justifiable without the direct application of Mill's canons, but only where background knowledge can play the role similar to a 'control group'.

In the absence of background knowledge, thematic affinity is insufficient evidence of a causal connection. But this is not yet a disagreement with Hopkins, since the method of

interpretation draws on background knowledge. Grünbaum must in fact hold a further condition, viz. that the background knowledge has been established by the (implicit or explicit) employment of Mill's canons. Unless this is a requirement, his insistence on Mill's canons as the only means to establish causal inferences is misplaced. So for Grünbaum to be right about commonsense psychology, it only needs to be the case that our general background knowledge regarding the typical connections between motives and human behavior is generated by a form of non-conscious statistical tracking that matches Mill's canons. Prima facie, this claim receives support from social psychology. Attribution theory provides an account of commonsense causal inferences, primarily ones relating to the social world. Kelley (1967; 1973) argued that we make inferences regarding the causes of people's behavior, including their motives and character traits, using criteria that roughly match Mill's Method of Difference. We aim to detect 'covariation' between cause and effect, and use three criteria to do so, viz. consistency, distinctiveness, and consensus. For example, to judge whether someone recommends a particular restaurant because it is good, we will be sensitive to whether they recommend it consistently; whether their recommendation is distinctive to that restaurant, or whether they tend to recommend everywhere they have eaten; and whether other people also recommend the restaurant. Kelley notes that in many cases of inference in which we do not have specific information regarding covariance, we draw on such information about similar causes and effects.

While supporting Grünbaum's view of commonsense psychological inferences, this account opens up the second line of reply to Grünbaum's attack on psychoanalysis noted in §1, viz. that in drawing inferences regarding the causes of neurosis, psychoanalysts employ Mill's canons implicitly. Non-conscious tracking of covariance across patients over time creates a significant body of background knowledge as to the causes of neurotic behavior and symptoms. Psychoanalysts can also contrast neurotic and 'normal' behavior, as the causes of the latter are

given by analysts' commonsense background knowledge concerning motives (which acts in place of a non-neurotic control group). However, this defence fails. First, the account of attribution theory just given is incomplete, and the full account (given below) turns out to support Hopkins' analysis and the method of interpretation. Second, this line of defense does not avoid the challenge, to be presented in §4, that our commonsense inferences – and so those of psychoanalysis – may be flawed.

§ 3.2 A conceptual analysis of desire

There are two reasons to think that Mill's Methods are insufficient to characterize commonsense psychological inference. The first objection derives from the logic of psychological concepts. Grünbaum (1993, 166) endorses an argument by Erwin (1992, 445-50) to the effect that we cannot, as Hopkins supposes, infer any causal explanations from the intentional content of a desire. That someone has a desire for a drink of water is no grounds for inferring that this desire is causally relevant to their getting a drink of water. The employment of Mill's Methods is necessary. The point can be read two ways: either that from the description of the desire, we have no grounds for inferring that it is in fact causally relevant to the action of getting a glass of water on this occasion; or that from the description of the desire, we cannot even infer the *kinds* of causal explanations in which it figures.

The second reading is clearly mistaken. It understands the existence and content of desires to be logically independent from behavior *in toto*. It allows the *possibility* that people in general may have desires which never have any influence on their behavior. (It is important to remember that desires may influence behavior in many more ways than merely leading one to seek their satisfaction, e.g. avowals, daydreams, etc.) But I can think of no (widely accepted) philosophical theory about the nature and content of desire that would not reject this possibility.

If we ask 'what *is* the desire to drink?', we can provide no answer without referring to the behavior to which it is typically causally relevant. What makes a mental state a *desire*, rather than some other type of state, is precisely the pattern of its relations (causal and/or normative) to behavior and other mental states. What gives a desire the *content* it has is also determined (in part) by such relations. What would make the desire to get a drink a *desire* and a desire *to get a drink* in the absence of any such relations? On the second reading of Erwin's claim, it appears a logical possibility that the desire to drink could, in fact, be *typically* causally relevant to people seeking sleep. 'Anything can cause anything', Hume taught us. But then nothing is left to the idea of it being a desire to get a drink.

For this reason, we cannot learn about motives and their effects in the first instance through the employment of Mill's Methods. To understand what a desire is, and what it is a desire for, is already to understand it in thematic affinity with the behavior to which it is typically causally relevant. What are we (or young children) inferring the existence *of*, exactly, if this connection is not already in place at the point of inference? The motive and behavior are not independent existences in such a way to allow Mill's canons to be the sole grounds for inferring a causal connection. The conceptual analysis of desire and content entail that we cannot and need not rely on Mill's Methods alone in the generation of background knowledge regarding the kinds of causal relations in which desires typically participate. (For many straightforward cases, as noted in the introduction, it may be better to say that we *recognize* rather than *infer* certain motives as expressed in behavior. But this is no objection to the line of argument here, as there is then no hard and fast line between recognition and inference.)

The first reading of Erwin's claim is that from the description of the desire, we have no grounds for inferring that it is in fact causally relevant to the action of getting a glass of water on this occasion. This is also wrong. Given the conceptual analysis just provided, we have a

defeasible, *prima facie* reason to infer from someone's getting a glass of water that she wants a glass of water. Of course, as discussed in §2.1, such an inference is not yet justified in the absence of further background evidence, and may be defeated by such. But this does not show that we must employ Mill's canons in securing the inference, as the next point indicates.

§3.3 Mill's canons and singular explanation

The second reason to think that Mill's canons are insufficient to characterize our inferences returns us to the inferences themselves. To justifiably attribute a particular motive on a particular occasion, we need 'background evidence that the act is not intended to serve some other purpose' (Erwin 1992, 448). For example, we cannot know that someone singing the national anthem is doing so because she wants to sing it (despite the thematic affinity), as she may be singing to curry favor with a patriot or to mock the audience. But how, using Mill's canons alone, are we to establish this background evidence, *specific to this occasion*, and so make the correct inference? Mill's canons establish relations between *types* of cause and their effects. Neither a statistical analysis of other occasions on which a singer has sung the national anthem, nor a statistical analysis of the various motives for which people generally tend to sing the national anthem, will alone provide an answer to why this singer sings it now. We go beyond Mill's canons in judging which possible cause is the actual cause in the specific situation, and for this, we are using further criteria for which motive provides the best explanation.

But what, then, of the earlier appeal to attribution theory? This was misleading through being incomplete. Attribution theory does not claim that judgments of 'covariation' are *all* we use, and it is clear that even our use of this only roughly corresponds to Mill's methods. For example, when Kelley (1972; 1973) discusses how we make inferences in situations in which we do not have the results of multiple observations, his remarks indicate that we employ a form of

inference to the best explanation. And developments in attribution theory since Kelley indicate that we take a host of factors into account and balance them against one another in a manner best described by inference to the best explanation (see Gawronski 2004). Furthermore, as we will see in §4.2, it appears that our observations and judgments of covariation are strongly influenced by cultural theories regarding the causes of behavior, and these theories prescribe criteria for 'best explanation' on which we draw in making our inferences.

It may be objected that this appeal to scientific psychology demonstrates that we are only able to establish the cogency of the method of interpretation by using Mill's canons. This is mistaken, however, as the two objections just considered, from conceptual analysis and from singular explanation, establish the method of interpretation independently. Furthermore, the appeal to scientific psychology is no objection to Hopkins. Grünbaum's position is that no causal *inference* can be justified except by Mill's canons. If Mill's canons show that some other non-Millian *method* can justify causal inferences, this forms an objection to Grünbaum, not Hopkins.

§ 3.4 The autonomy of interpretation

The two objections from conceptual analysis and singular explanation secure the conclusion that our commonsense psychological inferences cannot rest *solely* on the implicit employment of Mill's canons. However, they do not demonstrate Hopkins' stronger claim, that the method of interpretation is 'autonomous… and prior to such canons'. We have seen that the method of interpretation is logically and temporally prior to employing Mill's canons in fixing the relevant psychological concepts; but this activity does not exhaust the method of interpretation. Furthermore, social psychology indicates that we employ some roughly similar patterns of reasoning as part of our commonsense psychological inferences. We have arrived at an account that claims that we use a form of inference to the best explanation, within which something like

Mill's Method of Difference is embedded. This reflects the standard view that inference to the best explanation will be sensitive to precisely the sorts of covariations that Mill's methods also seek to identify (Rappaport (1996); Lipton (2004, 124f.)).

The issue of whether the method of interpretation is 'autonomous' is not one that needs settling for the argument to proceed. Grünbaum objected (§2.2) that psychoanalysts cannot use thematic affinity to support an inference from some starting point + free association to a mental state as cause of the starting point. Hopkins responds that the method of interpretation enables just this. I have argued that Grünbaum's objection to Hopkins' method of interpretation fails, as the method of interpretation forms an ineradicable part of our commonsense psychological causal inferences. Insofar as the method of interpretation can justify our everyday inferences, and psychoanalysis employs the same methodology, then Grünbaum's objection to causal inference in psychoanalysis fails.

However, Hopkins' defense of psychoanalysis depends on our commonsense psychological methodology being sound. Our attention, therefore, shifts from Grünbaum's formal concerns about methodology to more empirical ones. Social psychology has seen a sustained attack on the reliability of our usual practices in inferring people's motives, an attack that is independent of the question of the precise nature of our commonsense practices in relation to Mill's canons.[2]

§4. Inferring more than we know?

The concern for Hopkins' defense of psychoanalysis is that the psychoanalytic practice of inferring causes will inherit the flaws of commonsense practice, and so psychoanalysis cannot justify its causal claims using the methodology it shares with commonsense psychology. The challenge from social psychology needs to be met for Hopkins' defense to go through. The

social psychology debate over commonsense psychology ranges widely. For the purposes of discussion, I shall focus on issues arising out of the famous work of Ross, Nisbett and Wilson. I shall move from a more specific objection to a more general one.

§4.1 The fundamental attribution error and the psychoanalytic setting

The specific objection regards the 'fundamental attribution error' (Ross 1977). The fundamental attribution error has also sometimes been called the 'correspondence bias' (Gilbert and Jones 1986) – but, following Gawronski (2004), I distinguish them for reasons that will become clear. The correspondence bias is the tendency of observers to infer from a subject's behavior a corresponding disposition, even in situations in which the behavior is highly constrained. So, for example, in the original attitude attribution paradigm (Jones and Harris 1967), subjects are asked to read an essay defending a controversial moral position, and then asked whether the author holds the position argued for. In one condition, they are informed that the author had a free choice in which position to adopt. However, in the other condition, they are informed that the author was *instructed* to write an essay defending the position. They nevertheless tend to infer that the author holds the position defended.

The most common explanation for the correspondence bias refers back to the fundamental attribution error, viz. that observers hold the theory that behavior is caused by agents' dispositions rather than situational factors. Nisbett and Ross (1980, 31) claim that the 'most general and encompassing lay theory of human behaviour… is the assumption that behaviour is caused primarily by the enduring and consistent dispositions of the actor, as opposed to the particular characteristics of the situation to which the actor responds.' Much of the evidence for this comes from the correspondence bias, and supports the claim that commonsense psychological inferences can be unreliable.

Nisbett and Ross (1980, 244) charge Freud with 'elevating the fundamental attribution error to the status of a scientific principle' in assuming that all behavior can be traced to 'correspondent' motives, if only its 'real' meaning could be uncovered. This is a travesty of Freud's view, but the challenge is clear. Psychoanalysis has certainly emphasized, and possibly overemphasized, the causal role of motives and character traits in the explanation of behavior. And to employ the method of interpretation is to seek to connect behavior to motive through a connection in intentional content or 'meaning'.

This objection fails, however, because Nisbett and Ross are wrong about the fundamental attribution error. Despite its popularity, the evidence is against explaining the correspondence bias in terms of a fundamental attribution error. It is not that we hold a faulty *theory* of human behavior, as shown by the fact that the correspondence bias is corrected for when:

1. observers are in situations that encourage in-depth and careful processing of information, e.g. suspecting ulterior motives for the behavior;
2. observers have a high degree of what might be called 'social intelligence': individuals have varying degrees of complexity in their 'attributional schemata' which govern the causal inferences they make, as well varying degrees of motivation to explain human behavior;
3. observers have a high degree of 'negative capability', i.e. they do not have a strong need for explanatory closure, which forestalls their settling too quickly on an explanation in terms of the agent's attitudes. (Fletcher (1995, 73-9); Gawronski (2004))

On the basis of this evidence, we may argue that the correspondence bias supports only the conclusion that our causal inferences regarding people's dispositions can be unreliable in everyday situations that encourage casual or automatic responses. But the bias, and hence the unreliability, cannot be generalized to causal inferences made by individuals with complex attributional schemata, motivated to understand human behavior, working in situations that encourage in-depth and careful reflection, but still working with the methods and resources of commonsense psychology. These criteria clearly apply to the psychoanalytic clinical setting. The methodology of psychoanalysis does not rest on the fundamental attribution error, as Nisbett and Ross charge.

§4.2 Misattribution and heuristics

However, the fundamental attribution error is but one error. Nisbett and Wilson (1977) present a number of experiments demonstrating a series of errors in subjects' reports of the causes of their behavior (with corresponding errors for observers). Some examples:

1. *Insomnia*: Two groups of insomniacs were given a placebo pill. One group was told that it would produce symptoms of arousal, the other that it caused relaxation. As predicted, the 'arousal' group attributed their initial insomnia to the effects of the pill, and fell asleep sooner than usual (by 28%); the 'relaxed' group attributed their continued insomnia to being especially worried, and took longer to fall asleep (by 42%). When asked why they took more or less time to fall asleep, both groups produced rationalizations and rejected the experimenters' explanation in terms of the psychological effects of the placebo pill.

2. *Stockings*: Four pairs of stockings were placed on a table in a shopping centre and passers-by asked to identify which they preferred. Most indicated a preference, citing qualities such as smoothness, hue, etc. But there was a very strong positional effect, towards the right. Not only were the stockings identical, all but one subject rejected the suggestion that the position of the stockings influenced their preference.
3. *Drill*: Two groups of students were shown a film, and asked to rate their enjoyment of it. However, with one group, a power drill was used outside the room during part of the film; and they were asked to indicate whether this distraction had decreased their enjoyment. They indicated that it had. However, the ratings of the two groups were roughly equivalent – the drill had not decreased enjoyment, even though subjects thought that it had.

So we fail to recognize factors that are causally relevant, while we erroneously report factors that are not causally relevant. These mistakes, Nisbett and Wilson argue, are the effects of 'judgments about the extent to which a particular stimulus is a plausible cause of a given response', e.g. a difference in texture is a plausible cause of a preference, a difference in position are not. They suggest that judgments of plausibility and the causal explanations they support rest on cultural or, in some cases, idiosyncratic theories regarding cause-effect relations for behavior; or where these provide insufficient guidance, subjects generate new causal explanations 'by searching their networks of connotative relations surrounding the stimulus description and the response description' (1977, 248).

These claims reflect the famous work by Tversky and Kahneman on 'representativeness' and 'availability', heuristics that guide our inferences (Kahneman and Tversky (1972); Tversky and Kahneman (1973; 1974)). 'Representativeness' in causal inferences operates in our

judgments of whether a factor is a 'plausible' cause, whether it 'fits' with the effect (Nisbett and Ross 1980, 26-7). What counts as 'representative', in its simplest form, is governed by relations of similarity, e.g. complex effects require complex causes; but representativeness is also influenced by our cultural and idiosyncratic theories of typical cause-effect relations. Thematic affinity, we may infer from our previous discussions, is one aspect of representativeness, as we think a plausible cause of behavior with intentional content is a motive with the same, or related, content. Our inferences are also governed by the 'cognitive availability' of factors, which is a product of the initial strength or salience of the experience of the factor and the strength of verbal associations to it, among other things. Circumstances can reduce the availability of the actual cause, and enhance the availability of some non-causal factor, leading to a misattribution of causal relevance. Hence Nisbett and Wilson's conclusion that 'reports [on the causes of behavior] will be accurate when influential stimuli are (a) available and (b) plausible [i.e. representative] causes of the response, and when (c) few or no plausible but noninfluential factors are available' (1977, 253). These conditions apply equally to both subjects' reports and observers' inferences of the causes of a subject's behavior.

Precisely what should be inferred from this evidence? First, it must be emphasized (Nisbett and Ross 1980, 18, 23, 27) that no general rejection of commonsense methodology can be extrapolated. We cannot get from the reported experiments to a rejection of Hopkins' method of interpretation. First, the experimenters clearly rely on the kind of relations of thematic affinity Hopkins indicates, e.g. that verbal reports reflect conscious beliefs with the same content (preference for stockings, enjoyment of a film). In any case, that causal relations travel through thematic affinity is well-established experimentally, e.g. through manipulations using subliminal perception. And, furthermore, as Lynne Rudder Baker (1999; see also 1987, Ch. 7) has argued, the hypotheses of social psychology are couched in terms of commonsense

psychological concepts, and these bear a conceptual connection to typical cause-effect relations (as argued in §3.2). Social psychology therefore cannot support a wholesale rejection of the causal explanations offered in commonsense psychology. Second, 'the representativeness heuristic is a legitimate, indeed absolutely essential, cognitive tool. Countless inferential tasks… depend on deciding what class or category of event one is observing; such judgments inevitably hinge upon assessments of resemblance or representativeness' (Nisbett and Ross 1980, 27). In cases of classifying behavior and inferring motives, if Nisbett et al. are right, such assessments depend upon cultural 'theories' regarding cause-effect relations between motives and behavior, which therefore returns us directly to commonsense psychology (as cultural 'theory') and notions of thematic affinity (as embedded in commonsense psychology). The problem is not the representativeness heuristic itself, but the extent of its application.

On the issue of extent, then, Nisbett and Ross (1980, 211) note that the range over which we are accurate (where causally relevant factors are plausible and available, and we are not misled by noninfluential factors) is substantial – we are most often right. In such cases, the notion of 'representativeness', of which thematic affinity is one aspect, is a helpful guide. Furthermore, we may note that in many situations in which we fail to pick up on *all* causally relevant factors, we nevertheless are accurate about some. In the correspondence bias debate, for example, few argue that the attitudes of the agent are *irrelevant* to their behavior, even when situational factors are important causes. The case of *Insomnia* doesn't show that anxiety is not a factor in insomnia, while in *Drill*, it is hard to resist the prediction that many factors which we would expect to affect enjoyment (e.g. the quality of the film-making) affect enjoyment. Our identification of motivations and character traits requires a complementary corrective in terms of situational factors of which we are not conscious. Nevertheless, we reject factors that are actually relevant, e.g. position effects, and include factors that are not relevant at all (the drill). So the

evidence appears to support the conclusion that our everyday judgments of what is the 'best' explanation do not always lead to correct inferences – the methodology is flawed.

However, Fletcher (1995, 77) argues that we can generalize the response given in the case of the correspondence bias. Research in the area of people's causal inferences supports a *general* claim that 'there exists a patchwork of conditions under which biases or "errors" are reduced or eliminated'. The typical mistakes in causal inference identified by Nisbett and Wilson are more common in situations that encourage casual or automatic responses, and thoughtfulness, motivation, social intelligence etc. in general increases accuracy. More work is needed to support this claim further, but for now, we can note that, once again, these conditions apply in psychoanalysis.

A further consideration can be brought in support of psychoanalysis. Many of the causally relevant factors overlooked or rejected in Nisbett and Wilson's experiments are situational, rather than dispositional. However, the situation under which clinical data are produced remains stable and invariant as far as possible, so there are no varying situational factors to appeal to, to explain varying free associations and other behaviors. (Where there are changes in the situation, e.g. a rearrangement of the room or a forthcoming break in sessions, psychoanalysts are very alert to their potential impact.) Psychoanalytic methodology thus removes the potential for many of the errors Nisbett and Wilson identify. But might psychoanalysts falsely infer corresponding dispositions from behavior which is generated by constant situational factors within the clinical setting? Most psychoanalysts accept that the constant situational factors of the clinical setting alter analysand's behavior, e.g. lying on the couch may increase ability and/or willingness to talk freely. But such effects are either irrelevant to the inferences made or taken to facilitate the discovery of existing dispositions –

psychoanalysts do not, for instance, falsely conclude that their patients are more willing to discuss their problems than people who do not visit psychoanalysts!

How far such replies take us, we examine in the next, and last, section.

§ 5. Causal inference in psychoanalysis

§5.1 Recap

Hopkins' response to Grünbaum, we noted, had two elements:

> In commonsense psychological practice we already establish causal connections (in particular concerning the role of motives) interpretively, in ways that are autonomous, cogent, and prior to such canons. So it seems wrong to hold generally that cogency in a psychology of motive must satisfy them…
>
> Further, psychoanalytic theory seems an extension of commonsense understanding of motives, by interpretive means internal to it. (1988, 37)

We have assumed (§2.1) that Hopkins is correct about the second point. And we have agreed with Hopkins (§3.3), that a psychology of motive is, in part, established on the basis of thematic affinity, through the method of interpretation, though we have not established that this method is entirely autonomous with respect to Mill's canons, as it may, in part, rest on an implicit employment of the Method of Difference. The discussion of the debate in social psychology regarding the errors to which our causal inferences regarding behavior are prone raised the question of whether psychoanalysis inherits important inferential flaws from commonsense psychology. We have seen the beginnings of a response on behalf of psychoanalysis, viz. that further evidence in social psychology indicates that many of the errors to which we are prone

are reduced or eliminated in precisely the set of conditions that prevail in the clinical setting, including the social intelligence and motivation (and training) of the analyst. Developing this line of argument further could take us toward a defense of psychoanalysis as a form of expertise in inferring the motives of others within the clinical setting. This response disposes of the challenge from the fundamental attribution error, and mitigates the general challenge regarding the reliability of our commonsense psychological inferences.

§5.2 How far does the method of interpretation extend into psychoanalysis?

Nevertheless, we may be concerned that the reliability of the method of interpretation has only been demonstrated within commonsense psychology, and therefore its extension in psychoanalysis may take us outside the range within which reliance on this method (and the underlying representativeness and availability heuristics) can be justified. While some inferences using the method of interpretation may be justified, it does not follow that all are. What follows is only a brief sketch of an issue that requires more extensive consideration than I can give it here.

The psychoanalytic inferences Hopkins discusses are made from free association, together with its starting point, to a motive (or more commonly, motivating structure of mental states) as the cause of the clinical data. Just as such inferences in the case of commonsense psychology become both more justified and more explanatorily powerful as more evidence is accrued and more behavior explained, so it is with inferences in psychoanalysis, resting on the clinical data produced by the analysand over many sessions. However, there are many types or levels of inference in psychoanalysis:

1. regarding the specific motives for specific behaviors of an individual within an analysis;

2. regarding the typical structures of motivation, and their typical effects, that apply across many cases, e.g. the existence and nature of defense mechanisms, such as repression, projection, and so on, and their clinical manifestations;
3. regarding the causal role of these mental structures in the manifestations of neurosis and character traits;
4. regarding the causal origins of these structures, esp. in relation to childhood experience;
5. regarding larger scale structures within the mind, e.g. the superego in Freud, the depressive position in Klein, and their typical development in early childhood.

The first three of these types of inference correspond most closely with the reliable range of commonsense psychology, moving as they do from *current patterns of behavior to current motives and other dispositional structures*. The argument we have rehearsed therefore supports the claim that psychoanalytic clinical methodology justifies these forms of inference. Clinical data provide reliable evidence of the present existence of certain motivating and affective states. Psychoanalysts here extend familiar and established principles or patterns of causal inference to less familiar data, e.g. imagining something to be the case because one wishes it were (Hopkins 1988, §2) or irrational action stemming from unacknowledged desire or need (Miller 1988, 671). Furthermore, the motives and emotions inferred from the data are present, and can often be subjectively confirmed by the analysand, increasingly so as the analysis progresses. With this information also comes information about actual or potential psychic conflict. It is a central claim of psychoanalysis that such conflict is at the heart of neurosis, and the method of interpretation is capable of linking clinical data and such conflict in a way that can justify such a claim.

As specified so far, the inferences are from the clinical data to causal explanations *of the clinical data*. On many occasions, however, both analysand and analyst will try to account for the behavior of the analysand in everyday life, outside the clinical setting. Such inferences are justified to the extent that they appeal to structures of mental states already established on the basis of clinical evidence. As the method of interpretation makes clear, what is justified relates to a *pattern* – in this case, patterns of conflict. Such patterns within the clinical setting are the foundation for the claim that similar patterns of behavior outside the clinical setting (as reported by the analysand) manifest the same conflicts and are caused by the same constellation of psychological states. However, there may also be situational factors involved that neither analysand nor analyst are aware of, so the explanation may be incomplete and may overemphasize the causal role of dispositions. For example, unconscious motives may be overemphasized while non-motivated unconscious cognitive processes of the kind established by cognitive social psychology may be overlooked.

More theoretical inferences (types 4 and 5), such as the causal origins of neurosis, are not as well-supported by the argument we have made. Commonsense psychology moves from behavior to motives; but it has less to say on the origins of motives, and the soundness of the method of interpretation regarding the first is an insufficient reason to expect it to be sound regarding the second. We may anticipate, until it is otherwise shown, that there are too many factors of potential causal relevance to the formation of long-term dispositions, factors which may be either 'implausible' or 'unavailable'. This is not to say that clinical data will provide no evidence supporting theoretical inferences of this kind, but that clinical data alone may well not be enough to justify such inferences. Grünbaum is right, therefore, to think that from clinical data alone, we cannot reliably infer the original cause(s) of neurosis in general. Cooperative work

with other disciplines, such as neuroscience, developmental psychology and attachment theory, is necessary.

These conclusions are further supported by an appeal to the present reality of psychoanalysis. First, as psychoanalysis has developed, so greater emphasis has come to be placed on working in the transference in the present, and less emphasis has been given to a reconstruction of the origin of neurosis. The usual reason offered for this shift is that the clinical data cannot reliably support historical reconstructions, but can give insight into present conflict. Second, while different schools of psychoanalysis have developed a variety of theoretical accounts of neurosis and mental structure at a distance from the clinical phenomena, they are agreed on the existence and effect on behavior of unconscious defensive psychological processes (Miller 1988; Westen 1998, 334-5; Fonagy and Target 2003, Ch. 1.2). Third, many of the claims based on the first three types of inference are consilient with experimental data (Westen 1998), and the chances of a highly flawed method producing this kind of consilience are slim.[3]

References

Brearley, M. 2008. What do psychoanalysts do? In *The Academic Face of Psychoanalysis*, eds. L. Braddock and M. Lacewing, 20-32. London: Routledge.

Davidson, D. 1963. Actions, Reasons and Causes. *The Journal of Philosophy* 60, no. 23:685-700.

Erwin, E. 1992. Philosophers on Freudianism. In *Philosophical Problems of the Internal and External Worlds*, ed. J. Earman, 409-60. Pittsburgh: University of Pittsburgh Press.

Fletcher, G. 1995. *The scientific credibility of folk psychology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fonagy, P., and Target, M. 2003. *Psychoanalytic theories: Perspectives from developmental psychopathology*.

London: Routledge.

Freud, S. 1937. Constructions in analysis. In *The standard edition of the complete psychological works of Sigmund Freud*, ed., trans. J. Strachey, vol. 23: 255-269. London: Hogarth.

Gawronski, B. 2004. Theory-based bias correction in dispositional inference: The fundamental attribution error is dead, long live the correspondence bias. *European Review of Social Psychology* 15:183-217.

Gilbert, D., and Jones, E. 1986. Perceiver-induced constraint: Interpretations of self- generated reality. *Journal of Personality and Social Psychology* 50, no. 2:269-280.

Grünbaum, A. 1980. Epistemological Liabilities of the Clinical Appraisal of Psychoanalytic Theory. *Noûs* 14, no. 3:307-385.

Grünbaum, A. 1984. *The Foundations of Psychoanalysis*. Berkeley, CA: University of California Press.

Grünbaum, A. 1993. *Validation in the Clinical Theory of Psychoanalysis*. Madison, CT: International Universities Press, Inc.

Grünbaum, A. 2004. The Hermeneutic Versus the Scientific Conception of Psychoanalysis. In *Psychoanalysis at the Limit*, ed. J. Mills, New York: SUNY.

Hopkins, J. 1988. Epistemology and Depth Psychology. In *Mind, Psychoanalysis and Science*, eds. P. Clark and C. Wright, 33-60. Oxford: Blackwell.

Jones, E. 1990. *Interpersonal perception*. New York: Freeman.

Jones, E., and Harris, V. 1967. The attribution of attitudes. *Journal of Experimental Social Psychology* 3, no. 1:1-24.

Kahneman, D., and Tversky, A. 1972. Subjective probability: A judgement of representativeness. *Cognitive Psychology* 3, no. 3:430-454.

Kelley, H. 1967. Attribution theory in social psychology. In *Nebraska Symposium on Motivation*, ed. D. Levine, vol. 15:192-238. Lincoln: University of Nebraska Press.

Kelley, H. 1972. *Causal schemata and the attribution process*. New York: General Learning Press.

Kelley, H. 1973. The processes of causal attribution. *American Psychologist* 28, no. 2:107-128.

Kitcher, P. 1992. *Freud's Dream*. Cambridge, MA: MIT Press.

Levy, D. 1996. *Freud Among the Philosophers*. New Haven, CT: Yale University Press.

Manson, N. 2003. Freud's own blend. *Proceedings of the Aristotelian Society* 103:179-195.

Mill, J. S. 1904. *A System of Logic*. London: Longmans, Green & Co.

Miller, R. 1988. A Clinical Science. *Canadian Journal of Philosophy* 18, no. 4:659-679.

Nisbett, R., and Ross, L. 1980. *Human Inference: strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R., and Wilson, T. 1977. Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review* 84, no. 3:231-59.

Rappaport, S. 1996. Inference to the Best Explanation: Is it Really Different from Mill's Methods? *Philosophy of Science* 63, no. 1:65-80.

Ross, L. 1977. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In *Advances in experimental social psychology,* ed. L. Berkowitz, vol. 10: 173–220. New York: Academic Press.

Rudder Baker, L. 1987. *Saving Belief: a Critique of Physicalism*. Princeton, NJ: Princeton University Press.

Rudder Baker, L. 1999. What is this thing called commonsense psychology? *Philosophical Explorations* 2, no. 1:3-19.

Sachs, D. 1989. In Fairness to Freud: A Critical Notice of the Foundations of Psychoanalysis. *The Philosophical Review* 98, no. 3:349-378.

Sandis, C. ed. 2008. *New Essays in the Philosophy of Action*. Basingstoke: Palgrave-Macmillan.

Searle, J. 1983. *Intentionality*. Cambridge: Cambridge University Press.

Tversky, A., and Kahneman, D. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5, no. 2:207-232.

Tversky, A., and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185:1124-1131.

Westen, D. 1998. The Scientific Legacy of Sigmund Freud. *Psychological Bulletin* 124, 333-71.

Wilson, T. 2002. *Strangers to Ourselves*. Cambridge, MA: Belknap Press.

[1] A number of authors (Levy 1996, ???; Sachs 1989, 373; Miller 1988, 658; Hopkins 1988) point out that in certain passages (e.g. 1984, 200) Grünbaum, perhaps inconsistently, allows precisely the kind of causal inference Hopkins defends. For example, from the slip, said by a man at a party to a woman with a large bust and low-cut dress, 'Excuse me, I must slip out for a breast of flesh air', Grünbaum allows us to infer that the slip was caused by a distracting thought about the woman's breasts. But in this passage (and others), Grünbaum has not given the *methodology* by which such an inference is justified, only the grounds. Consistent with the arguments presented in the next section, I expect Grünbaum would argue that we employ significant background knowledge in making and justifying the inference. He has not, therefore, allowed the use of thematic affinity as a ground for inference even in such, so-called 'transparent' cases.

[2] Like Hopkins, Miller (1988) argues that the methodology for justifying causal inferences in psychoanalysis is a form of inference to the best explanation – the justification rests in 'inferences from mutually supporting successes in explaining what needs explaining' (672). However, he argues that inference to the best explanation, rather than Mill's canons, is the *normal* scientific methodology for establishing causes (672). There is therefore nothing unusual or especially distinctive in the methodology of psychoanalysis that requires separate discussion (e.g. the method of interpretation). But it is possible to make too much of this claim. Not all 'inference to the best explanation' may operate in the same way, and normal scientific methodology usually utilizes *data generated by the explicit application of Mill's methods* while it is clear that clinical methodology does not. Furthermore, Miller faces the same challenge as Hopkins, viz. to find a response to concerns that the ways in which psychoanalytic inferences resemble commonsense psychological ones mean that psychoanalysis equally inherits the flaws on commonsense psychology. His conclusion is not dissimilar to that defended in §5. Miller claims that a number of traditional psychoanalytic causal claims, including the origins of neurosis in childhood events, are unjustified, and that his defense extends to a more limited range of claims, which are held to comprise the 'core' contemporary psychoanalytic commitments (659). However, while he indicates that this restriction in necessary on grounds of the clinical evidence since Freud (662), it is unclear whether he thinks the restriction is also necessary on methodological grounds, i.e. whether the specific causal principles that underpin psychoanalytic inferences *could* support the more ambitious traditional claims, or whether their reach is limited to the core commitments he outlines.

[3] Thanks to Louise Braddock, Richard Gipps, Constantine Sandis, Thad Metz, Katie Page, Stacie Friend, Tom Crowther, and the members of seminars in St John's College, Oxford, Rhodes University, and Heythrop College for their many helpful comments and references. A very special thanks to my research assistant, Maarten Steenhagen, for all the help he provided, and to Heythrop College, for the grant that made his assistance possible.